Homework №5

On the convexity of the loss function of logistic regression
(special case: binary classification problem)

Due: Friday, 17.04.2020
Assignment weight: 0.2
Nominal max. score: 100pts.
Effective max. score: 20pts.

RECAP:

Logistic regression is a special case of Generalized Linear Models, where the parameter $p(\theta, x_i)$ of a Bernoullli distribution may be interpreted as a probability for the object x_i to be of a class "1". $p(\theta, x_i)$ is given by:

$$p(\theta, x_i) = \frac{1}{1 + \exp(-\theta \cdot x_i)}$$

Consider the classes encoding such as $y_i = 0$ for the "negative" examples, and $y_i = 1$ for the "positive" examples. Then the likelihood of the sample $\mathcal{T} = \{y_i, x_i\}\,, i = 1 ... N$ is given by the following:

$$L(\mathcal{T}, \theta) = \prod_{i=1}^{N} \left(p_i^{y_i} * (1 - p_i)^{(1-y_i)}\right)$$

Logarithm of the likelihood is given by:

$$\begin{aligned} \ell(\mathcal{T}, \theta) &= \sum_{i=1}^{N} \left(y_i * \ln p_i + (1 - y_i) * \ln\left(1 - p_i\right)\right) = \\ &= -\sum_i \ln\left(1 + \exp\left(\theta \cdot x_i\right)\right) + \sum_i y_i * \theta \cdot x_i \end{aligned}$$

According to the Maximum Likelihood (ML) principle, the loss fuunction here is negative log. likelihood:

$$\mathcal{L}(\mathcal{T}, \theta) = -\ell(\mathcal{T}, \theta)$$

Please refer to the Lecture 11 and mentioned additional reading materials for more information.

Assignment:

Prove that the loss function $\mathcal{L}(\mathcal{T}, \theta)$ is convex everywhere in Θ space.